Extension of 3DVAR to 4DVAR: Implementation of 4DVAR at the Meteorological Service of Canada

PIERRE GAUTHIER, MONIQUE TANGUAY, STÉPHANE LAROCHE, AND SIMON PELLERIN

Meteorological Research Division, Environment Canada, Dorval, Québec, Canada

JOSÉE MORNEAU

Meteorological Service of Canada, Environment Canada, Dorval, Québec, Canada

(Manuscript received 19 April 2006, in final form 29 August 2006)

ABSTRACT

On 15 March 2005, the Meteorological Service of Canada (MSC) proceeded to the implementation of a four-dimensional variational data assimilation (4DVAR) system, which led to significant improvements in the quality of global forecasts. This paper describes the different elements of MSC's 4DVAR assimilation system, discusses some issues encountered during the development, and reports on the overall results from the 4DVAR implementation tests. The 4DVAR system adopted an incremental approach with two outer iterations. The simplified model used in the minimization has a horizontal resolution of 170 km and its simplified physics includes vertical diffusion, surface drag, orographic blocking, stratiform condensation, and convection. One important element of the design is its modularity, which has permitted continued progress on the three-dimensional variational data assimilation (3DVAR) component (e.g., addition of new observation types) and the model (e.g., computational and numerical changes). This paper discusses some numerical problems that occur in the vicinity of the Poles where the semi-Lagrangian scheme becomes unstable when there is a simultaneous occurrence of converging meridians and strong wind gradients. These could be removed by filtering the winds in the zonal direction *before* they are used to estimate the upstream position in the semi-Lagrangian scheme. The results show improvements in all aspects of the forecasts over all regions. The impact is particularly significant in the Southern Hemisphere where 4DVAR is able to extract more information from satellite data. In the Northern Hemisphere, 4DVAR accepts more asynoptic data, in particular coming from profilers and aircrafts. The impact noted is also positive and the short-term forecasts are particularly improved over the west coast of North America. Finally, the dynamical consistency of the 4DVAR global analyses leads to a significant impact on regional forecasts. Experimentation has shown that regional forecasts initiated directly from a 4DVAR global analysis are improved with respect to the regional forecasts resulting from the regional 3DVAR analysis.

1. Introduction

Over the last few years, the variational form of statistical estimation has been implemented at many operational centers. The motivation originated from the difficulties associated with the assimilation of satellite data such as the Television Infrared Observational Satellite (TIROS-N) Operational Vertical Sounders (TOVS) radiances. Lorenc (1986) showed that the statistical estimation problem could be cast in a variational form [the three-dimensional variational data assimilation (3DVAR)], which is just a different way of solving the problem that the so-called *optimal interpolation* attempts to solve directly. Eyre (1989) showed, in a 1DVAR context, that a variational formulation leads to a more natural framework for the direct assimilation of radiances instead of *retrieved* temperature and humidity profiles. This is also true for any indirect measurement of the state of the atmosphere. Talagrand and Courtier (1987) showed that the use of the adjoint of a numerical model makes it possible to determine the initial conditions leading to a forecast that would best

Corresponding author address: Dr. Pierre Gauthier, Data Assimilation and Satellite Meteorology Section, Meteorological Research Division, Environment Canada, 2121 Trans-Canada Highway, Dorval, QC H9P 1J3, Canada.
E-mail: pierre.gauthier@ec.gc.ca

DOI: 10.1175/MWR3394.1

fit data available over a finite time interval. These two formulations can be combined to yield what is now called the four-dimensional variational data assimilation (4DVAR) formulation of the statistical estimation problem. This approach has been used operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) since 1997 (Rabier et al. 2000) and at Météo-France since 2000 (Gauthier and Thépaut 2001). Other centers have also recently implemented a 4DVAR global assimilation system (e.g., the Met Office and the Japan Meteorological Agency).

At the Meteorological Service of Canada (MSC), the first implementation of a 3DVAR system occurred in 1997 (Gauthier et al. 1999a). The system was then designed to be as close as possible to the previous optimal interpolation system. Since then, substantial modifications have been brought to different aspects of the system. The 3DVAR was reformulated to define the analysis increments on the model's own vertical coordinate and a new formulation of the background error covariances was introduced (Gauthier et al. 1999b). This framework was more appropriate to add new data in the system [e.g., Advanced TOVS (ATOVS) radiance data; Chouinard et al. 2001]. A variational quality control was also included (Ingleby and Lorenc 1993; Andersson and Järvinen 1999; Gauthier et al. 2003).

This paper presents the strategy used to couple the 3DVAR and the Global Environmental Multiscale (GEM) model to obtain an incremental 4DVAR assimilation system in which the model and the 3DVAR components are kept as separate entities. The 3DVAR code includes the observation operators and all that is required to produce a 3DVAR analysis (e.g., background error covariances). Then the extension to 4DVAR only requires two more operators: the tangent linear and adjoint of the forecast model. The approach retained was then to use coupling mechanisms to link the model code to that of the 3DVAR. The main benefit of this approach is to considerably simplify the development and maintenance of the codes of the modeling assimilation system. On the one hand, work was going on to add new data types within the 3DVAR while the tangent-linear and adjoint models were developed and tested. The development of the tangent-linear and adjoint models has been developed and these were first used to perform sensitivity studies (Laroche et al. 2002b). The 4DVAR was finally implemented in the operational suite in March 2005.

The paper is organized as follows. Section 2 presents the formulation of the variational assimilation first in its 3D version and then its extension to 4D. The extension of a 3DVAR to 4DVAR is presented in section 3 with some preliminary validation tests and experiments. Section 4 discusses some problems associated with the numerics of the tangent-linear and adjoint model formulations that had to be addressed. Section 5 discusses the choices made for the final configuration of the 4DVAR. This includes the characteristics of the simplified model used in the incremental 4DVAR and in particular, the impact of the simplified physics. The role of the outer iterations is also discussed. Section 6 presents an overview of the impact of 4DVAR during the two-month assimilation cycles for summer and winter. The results clearly show that 4DVAR makes better use of the satellite data than 3DVAR. In a companion paper, Laroche et al. (2007) present a more complete analysis and diagnostics of a complete suite of experiments performed in preparation for the implementation.

2. Incremental formulation of variational assimilation

The statistical estimation problem can be cast in a variational form as

$$J(\mathbf{X}) = \frac{1}{2}(\mathbf{X} - \mathbf{X}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{X} - \mathbf{X}_b) + \frac{1}{2}[\mathbf{H}(\mathbf{X}) - \mathbf{y}]^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}(\mathbf{X}) - \mathbf{y}] \equiv J_b(\mathbf{X}) + J_o(\mathbf{X}), \quad (1)$$

where $\mathbf{B}$ is the background error covariance matrix, $\mathbf{X} = [\mathbf{u}^{\mathrm{T}}, \mathbf{v}^{\mathrm{T}}, \mathbf{T}^{\mathrm{T}}, (\ln \mathbf{q})^{\mathrm{T}}, \mathbf{p}_s^{\mathrm{T}}]^{\mathrm{T}}$ stands for the state vector (of dimension N), $\mathbf{X}_\mathrm{b}$ is the background state, $\mathbf{H}$ is the nonlinear observation operator, $\mathbf{y}$ is the data vector (of dimension M), and $\mathbf{R}$ is the observation error covariance matrix. Gauthier et al. (1999a) give a detailed description of the first implementation of 3DVAR at MSC. In its incremental form (Courtier et al. 1994), the variable $\Delta\mathbf{X}^{(k)} = \mathbf{X} - \mathbf{X}^{(k)}$ is introduced as a departure from a reference state $\mathbf{X}^{(k)}$ and the nonlinear operator $\mathbf{H}$ is approximated as

$$\mathbf{H}(\mathbf{X}) \equiv \mathbf{H}[\mathbf{X}^{(k)}] + \mathbf{H}'^{(k)}\Delta\mathbf{X}^{(k)}, \quad (2)$$

where $\mathbf{H}'^{(k)} = \partial\mathbf{H}/\partial\mathbf{X}$ is the Jacobian of the nonlinear observation operator evaluated at $\mathbf{X}^{(k)}$: this is referred to as the *tangent linear* of $\mathbf{H}$. Provided $\Delta\mathbf{X}$ is sufficiently small, (1) can be approximated as

$$J^{(k)}[\Delta\mathbf{X}^{(k)}] = \frac{1}{2}[\mathbf{X}^{(k)} - \mathbf{X}_b^{(k)}]^{\mathrm{T}}\mathbf{B}^{-1}[\Delta\mathbf{X}^{(k)} + \Delta\mathbf{X}_b^{(k)}] + \frac{1}{2}[\mathbf{H}'^{(k)}\Delta\mathbf{X}^{(k)} - \mathbf{y}']^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}'^{(k)}\Delta\mathbf{X}^{(k)} - \mathbf{y}'], \quad (3)$$

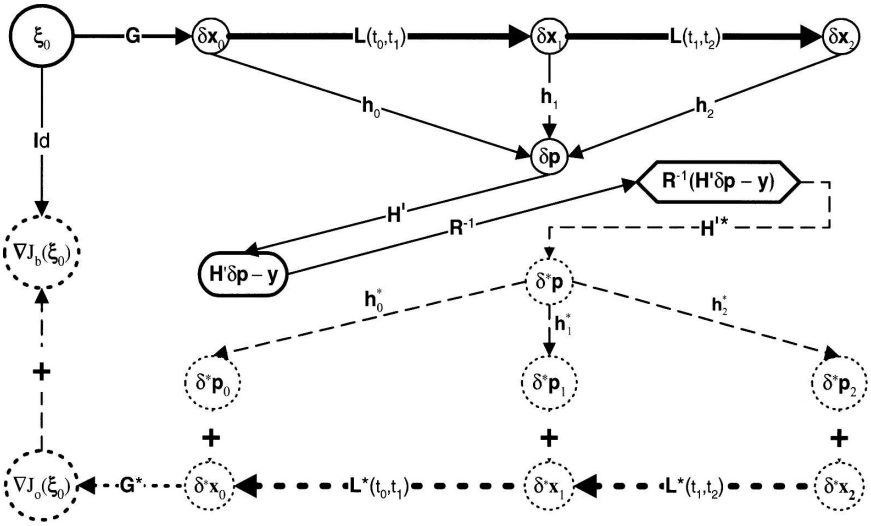

FIG. 1. Schematic representation of the data flow within a single iteration of 4DVAR. Thin and thick lines are associated with operators included within the 3DVAR and the GEM model, respectively. Dashed lines indicate the adjoint of operators. Here $\delta\mathbf{x}_k \equiv \delta\mathbf{x}(t_k)$.

where $\Delta\mathbf{X}_b^{(k)} \equiv \mathbf{X}^{(k)} - \mathbf{X}_b$ is the departure of the reference state $\mathbf{X}^{(k)}$ from the background state, while $\mathbf{y}' \equiv \mathbf{y} - \mathbf{H}[\mathbf{X}^{(k)}]$ represents the observation departure from the same reference state. In the incremental approach, the full nonlinear problem can be solved by minimizing (3) provided the reference state is updated regularly (Laroche and Gauthier 1998). If $\mathbf{H}$ is linear, (3) is then strictly equivalent to (1).

In 3DVAR, the resolution of the analysis increment is dictated by the nature of the structure functions introduced in the background error covariances, in which case (3) can be solved at a lower resolution (Laroche et al. 1999; Gauthier 2003). Introducing $\delta\mathbf{x}^{(k)} \equiv \mathbf{\Pi}_L\Delta\mathbf{X}^{(k)}$ in (3) leads to

$$J_L^{(k)}[\delta\mathbf{x}^{(k)}] = \frac{1}{2}[\delta\mathbf{x}^{(k)} + \mathbf{\Pi}_L\Delta\mathbf{X}_b^{(k)}]^{\mathrm{T}}\mathbf{B}^{-1}[\delta\mathbf{x}^{(k)} + \mathbf{\Pi}_L\Delta\mathbf{X}_b^{(k)}] + \frac{1}{2}(\mathbf{H}'^{(k)}\delta\mathbf{x}^{(k)} - \mathbf{y}')^{\mathrm{T}}\mathbf{R}^{-1}[\mathbf{H}'^{(k)}\delta\mathbf{x}^{(k)} - \mathbf{y}'], \tag{4}$$

where $\mathbf{\Pi}_L$ is the interpolation from the high- to low-resolution grid, $\mathbf{\Pi}_L^{-I}$ is its *generalized inverse* and it corresponds to the interpolation of the low-resolution increment to a high-resolution grid.[1] In this context, the analysis will only correct the large-scale components of the background state. Finally, if $\mathbf{B} = \mathbf{B}^{1/2}\,\mathbf{B}^{1/2\mathrm{T}}$, we let

$$\boldsymbol{\xi}^{(k)} = \mathbf{B}^{-1/2}\delta\mathbf{x}^{(k)}$$

$$\bar{\boldsymbol{\xi}}_k = \mathbf{B}^{-1/2}\mathbf{\Pi}_L(\mathbf{X}^{(k)} - \mathbf{X}_b)$$

to get

$$J_L^{(k)}(\boldsymbol{\xi}^{(k)}) = \frac{1}{2}(\boldsymbol{\xi}^{(k)} + \bar{\boldsymbol{\xi}}_k)^{\mathrm{T}}(\boldsymbol{\xi}^{(k)} + \bar{\boldsymbol{\xi}}_k) + \frac{1}{2}(\mathbf{H}'^{(k)}\mathbf{G}\boldsymbol{\xi}^{(k)} - \mathbf{y}')^{\mathrm{T}}\mathbf{R}^{-1}(\mathbf{H}'^{(k)}\mathbf{G}\boldsymbol{\xi}^{(k)} - \mathbf{y}'), \tag{5}$$

with $\mathbf{G} = \mathbf{B}^{1/2}$. The whole minimization procedure then consists of minimizing (5) by performing a certain number of *inner* iterations. The resulting low-resolution increment is then added to the high-resolution reference state through an *outer* iteration that updates the reference state to

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \mathbf{\Pi}_L^{-I}\mathbf{G}\boldsymbol{\xi}_*^{(k)}. \tag{6}$$

Given that $\mathbf{\Pi}_L\mathbf{\Pi}_L^{-I} = \mathrm{I}$, it then follows that $\bar{\boldsymbol{\xi}}_{k+1} = \boldsymbol{\xi}_*^{(k)} + \bar{\boldsymbol{\xi}}_k$, identically, with $\boldsymbol{\xi}_*^{(k)}$ being the convergence point of the previous inner minimization. The role of the outer iteration is to update the tangent-linear definition of the observation operator and to reevaluate the observation departures $\mathbf{y}'$.

As proposed in Courtier (1997), extending 3DVAR to 4DVAR is best understood by including the model integration as part of the observation operator. So, if observations are available within the time interval $t_0 \le t < t_K$, the observation operator for any observation at time t_j is $\mathbf{H}_j'^{(k)}\mathbf{L}(t_0, t_j)$, where $\mathbf{L}(t_0, t_j)$ is the propagator of the tangent-linear model (TLM) obtained from

[1] This operation can be viewed as a generalized inverse representing the interpolation of a low-resolution increment on the high-resolution grid. If this is followed by Π_L, this should recover exactly the low-resolution increment. This is true if this operation is done in spectral space for instance.

TABLE 1. Description of the different spaces involved in the description of the assimilation problem.

Object	Definition
$\boldsymbol{\xi}$	Canonical space
$\delta\mathbf{X}(t_k) \equiv \delta\mathbf{X}_k$	Model-state perturbation in physical space at time t_k comprising the horizontal wind components, u and v; temperature, T; logarithm of specific humidity, ln q; and surface pressure, p_s.
$\delta\mathbf{p}_k$	Profiles of model state at observation locations and time t_k
$\mathbf{w}$	Observation space variable

a linearization around a trajectory of the nonlinear model. Figure 1 gives a schematic of the operations involved in a single inner iteration of 4DVAR, including the adjoint of the operators required for the evaluation of the gradient of the observation component of the cost function. One immediately sees that it is possible to build a 4DVAR by using coupling mechanisms through which the 3DVAR component can make requests to the model to obtain a sequence of model states at observation times in response to a perturbation $\delta\mathbf{x}_0$ defined at the initial time. This is how the 4DVAR has been built at MSC. More details are given in the next section. For the time being, it is important to stress that, at each outer iteration, the linearization of both the observation operators $\mathbf{H}'$ and the TLM (and its adjoint) are updated.

3. The modular character of assimilation algorithms

From the discussion of the previous section, the algorithm is made up of basic *units* that exchange different types of *objects*. This idea is exploited in Lagarde et al. (2001) to show that by building these basic units, it is then feasible to obtain a wide range of assimilation algorithms. Table 1 describes in general terms the nature of the objects involved while Table 2 gives a short summary of the operators needed within a single inner iteration of the minimization. Referring to Fig. 1, the extension from 3DVAR to 4DVAR can be achieved by introducing a coupling between the *3DVAR* module, which includes all operations associated with the background error covariances and the observation operators: this module will be referred to as the background and observations (BGOBS) from now on to avoid confusion with 3DVAR. On the other hand, the *model* module includes the tangent-linear and adjoint models.[2] In the direct branch, the BGOBS module provides the initial conditions to the model module that can then perform its integration from time t_0 to time t_L. Prior to this, during the initialization phase, the time and location of each observation has been communicated to the model so that it only has to output profiled model states $\delta\mathbf{p}_k$ at the observation locations, which are then sent back to BGOBS. These profiles are obtained from a cubic horizontal interpolation performed within the model. At this point, BGOBS can initiate the same treatment as in a 3DVAR analysis because the observation operators only require profiled model states as input. It should be said that the tangent linear of the observation operators is defined here with respect to a profiled model state obtained from the nonlinear model with its complete set of parameterizations. Similarly, the adjoint of the observation operators yields profiled model states that are sent back to the model. The adjoint of the horizontal interpolation is performed first at the observation time to define the input needed for the backward integration of the adjoint model. For one iteration, the cost of two model 6-h integrations (TLM and adjoints) takes 61% of the time while the BGOBS component takes only 30%. There are only two exchanges of data between the model and BGOBS and

[2] This approach raises scheduling issues with the operating system when the same computing resources are shared between two simultaneously running units. For instance, once the BGOBS module has sent information to the model module, it immediately puts itself into I/O wait mode and its resources become available to the model as the two modules never have to run simultaneously. This approach was tested first on the NEC SX-6 computer and then on the IBM-p690. More details on this can be found in Pellerin et al. (2005).

TABLE 2. Summary of the basic operators required to perform an inner iteration of the minimization.

Direct operators			Adjoint (*) operators		
Input	Operator	Output	Input	Operator	Output
$\boldsymbol{\xi}$	$\mathbf{G} = \mathbf{B}^{1/2}$	$\delta\mathbf{X}_0$	$\delta\mathbf{X}_0^*$	$\mathbf{G}^* = \mathbf{B}^{1/2\mathrm{T}}$	$\boldsymbol{\xi}^*$
$\delta\mathbf{X}_k$	$\mathbf{L}(t_k, t_{k+L})$ tangent-linear integration from time t_k to t_{k+L}	$\delta\mathbf{X}_{k+L}$	$\delta\mathbf{X}_{k+L}^*$	$\mathbf{L}^*(t_k, t_{k+L})$ backward adjoint model integration from time t_{k+L} to t_k	$\delta\mathbf{X}_k^*$
$\delta\mathbf{X}_k$	$\mathbf{h}$: Horizontal interpolation	$\delta\mathbf{p}_k$	$\delta\mathbf{p}_k^*$	$\mathbf{h}^*$: Adjoint of the horizontal interpolation	$\delta\mathbf{X}_k^*$
$\delta\mathbf{p}_k$	$\mathbf{H}$ (or $\mathbf{H}'$): Observation operator (or its *tangent linear*)	$\mathbf{w}_k$	$\mathbf{w}_k^*$	$\mathbf{H}'^*$: Adjoint of the tangent linear of the observation operator	$\delta\mathbf{p}_k^*$

this takes a negligible amount of time (7%). For a more detailed discussion of our Multiple Program Multiple Data (MPMD) implementation, the reader is referred to Pellerin et al. (2005).

The profiled states offer the advantage of being grid independent except for the vertical levels that depend on the vertical coordinate used. However, the observation operators have all been coded such that they only require the knowledge of the pressure values at each level: no assumptions are made about these values. Details are given in the appendix.

As the nonseparable homogeneous and horizontally isotropic correlations are represented in spectral space as a block-diagonal matrix (Gauthier et al. 1999b), BGOBS uses a Gaussian grid in the global analysis, which is convenient for spectral transforms. The extension to 4DVAR required that the increment be interpolated to the model's own grid, which differs from the Gaussian grid. Moreover, the model's variables do not correspond to the physical variables as defined in 3DVAR and an additional change of variables is needed. For example, the model uses specific humidity and not its logarithm: these changes are introduced in the model altogether with their tangent-linear and adjoint versions.

The schematic of 4DVAR shown in Fig. 1 indicates that if the TLM integration is replaced by the identity, then 4DVAR is strictly equivalent to 3DVAR. This test has been used to validate the coupling between the model and BGOBS. In particular, it was possible to determine the impact of a significant change in accuracy because the GEM model is hard coded in single precision accuracy. From the discussion in section 2, several data exchanges between BGOBS and the model take place in the evaluation of the 4DVAR cost function. First, the BGOBS unit sends the analysis increment defined in physical space to GEM. Second, GEM integrates over the assimilation period and interpolates the vertical profiles of analysis increments at prescribed observation times and locations. Finally, GEM sends back those profiles to BGOBS to complete the evaluation of the cost function. The adjoint of those operations are needed to compute the gradient. To validate those operations, experiments were designed in which all operations are activated except for the model integration. This way, the results of a 3DVAR should be recovered, or at least make little difference. Two elements make this process not entirely transparent. The first is linked to the conversion of variables. BGOBS deals with temperature, logarithm of specific humidity, surface pressure, and wind components whereas the associated control variables of GEM are virtual temperature, specific humidity, logarithm of a scaled surface pressure, and image winds (i.e., wind components multiplied by the cosine of the latitude). The second difference is associated with the grids. As mentioned earlier, BGOBS keeps the analysis increment on a Gaussian grid due to the treatment of the background term J_b with its associated spectral calculations whereas GEM uses an Arakawa-C grid in the horizontal with a uniform resolution (see Fig. 5 in Côté et al. 1998a).

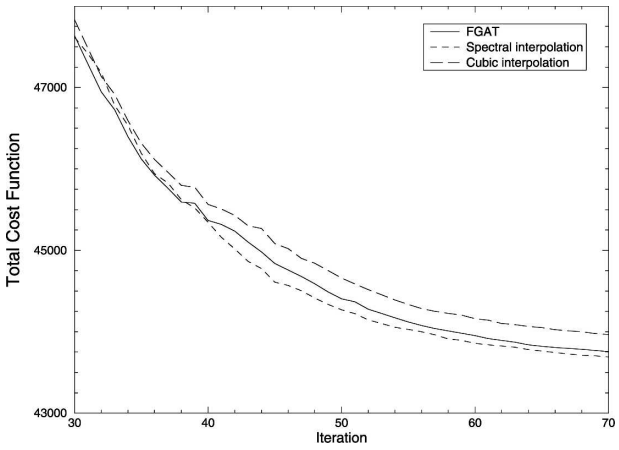

FIG. 2. Representation of the cost function in 4D-Identity experiment compared to a 3DVAR [first guess (background field) at the appropriate time (FGAT)] when spectral or cubic interpolation to the Arakawa-C staggered grid is used.

The first change was to use a Gaussian grid in GEM for the scalar fields, but the staggered winds need to be obtained from a horizontal interpolation. Figure 2 shows the resulting cost function when cubic interpolation is used to place the winds on the staggered grid and when a spectral interpolation is used to go directly from the spectral representation in terms of vorticity and divergence to the winds on the staggered grid. In these experiments, the full coverage of observations is used. To emulate what is done in the 3DVAR, the profiled model states at the observation locations are obtained here with a linear interpolation by using the grid of the requested variable.[3] The results show that the passage to the staggered grid with cubic interpolation causes the minimization to saturate at a level higher than what comes out from 3DVAR. Using the spectral interpolation manages to validate the 4D-Identity against the results of 3DVAR. Although close, small differences still remain. Those come from the change of physical variables described earlier and also from the lower numerical accuracy used to represent model fields in the GEM model.

[3] The operational implementation and all experiments presented later in this paper use cubic interpolation to obtain the profiled model states.

4. The tangent linear and adjoint of a gridpoint semi-Lagrangian model

The model is the Canadian GEM model (Côté et al. 1998a,b), currently used in operations at the MSC. The model operates on a 400 × 200 horizontal latitude–longitude grid with 28 vertical levels up to 10 hPa in a terrain-following coordinate. A more detailed description can be found in Côté et al. (1998b). It was recently redesigned for a parallel distributed-memory implementation with explicit message passing (Qaddouri et al. 1999; Desgagné et al. 2000). The linearization of the model and its transposition were done by hand. Numerics of the model and their impact on the tangent-linear and adjoint counterparts were studied. Polavarapu et al. (1996) and Tanguay et al. (1997) studied the linearization of the process of interpolation in the semi-Lagrangian advection scheme while Polavarapu and Tanguay (1998) examined the linearization of the iterative processes involved in the semi-Lagrangian scheme and in the solver. Finally, Tanguay and Polavarapu (1999) presented a complete analysis of the transposition of the semi-Lagrangian passive tracer equation. The linearization of a simplified physics package containing parameterizations of vertical diffusion (Laroche et al. 2002a), subgrid-scale orographic drag, and stratiform precipitation (Zadra et al. 2004) was developed and validated.

The GEM model uses a semi-Lagrangian two-time-level scheme that can become numerically unstable (Pudykiewicz and Staniforth 1984; Smolarkiewicz and Pudykiewicz 1992; Durran 1999). The first applications of the TLM and its adjoint have been for a posteriori sensitivity studies in which a correction to the initial conditions is sought to correct the forecast error at a 24-h lead time (Laroche et al. 2002b). Those studies managed to identify some deficiencies in the model in the vicinity of the polar regions, particularly the Antarctic, and near the surface. The problems in the polar regions were associated with the increased horizontal resolution resulting from the convergence of the meridians in a gridpoint model. Combined with strong winds, this can lead to situations where the two-time-level semi-Lagrangian scheme becomes unstable by violating the *Lipschitz condition:*

$$L_S = \left| \Delta t \frac{\Delta u}{\Delta \lambda} \right| < 2,$$

L_S being the Lipschitz number. This condition is required to insure convergence of the iterative process involved in locating the upstream position. Polavarapu and Tanguay (1998) showed that this creates problems in the TLM for the linearized version of this iterative process. The problem impacts the convergence of the 4DVAR minimization itself, which failed to converge in a situation where the Lipschitz criterion was not met. This problem is more acute in the polar regions and, in a shallow-water model, can be controlled by increasing locally the horizontal diffusion, applied at the end of the time step (Tanguay and Polavarapu 1999). This manages to damp the spurious small scales emerging due to the convergence of the meridians. However, in the baroclinic case, this is not sufficient.

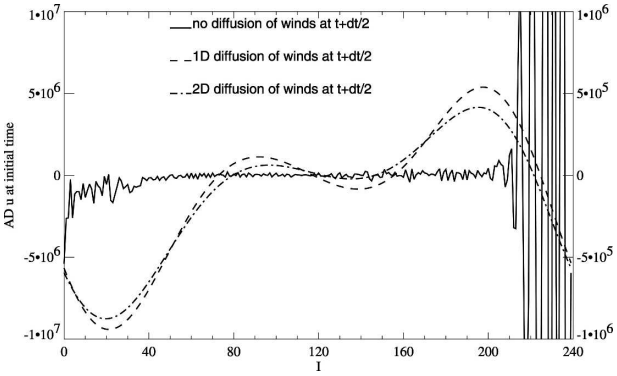

FIG. 3. Impact of horizontal diffusion on adjoint model integrations over 6 h. The curves represent an adjoint integration without any diffusion (solid line), with the full 2D diffusion (dot–dashed line), and the 1D diffusion scheme (dashed line) used in the 4DVAR. These results are shown on the closest latitude circle surrounding the South Pole near the surface.

A reexamination of the two-time-level scheme showed that it is necessary to apply horizontal diffusion on the wind field *before* they are used to estimate the upstream position. This has to be done twice in our two-time-level scheme that employs a Crank–Nicholson process to estimate the winds at the halftime step (Côté et al. 1998a). Predictor values are obtained for the dynamical variables and those are used to compute the different terms associated with the contributions from the physical parameterizations. As in the nonlinear model, horizontal diffusion could be applied on those predictor dynamical variables before they are used within the physical parameterizations. Figure 3 shows the impact of these changes on the resulting adjoint integration over a period of 6 h. The new scheme leads to a significant reduction in the local value of the Lipschitz number (not shown). Near the Poles, the reduction in grid length occurs only in the zonal direction. The purpose of the added diffusion is to filter out those spurious small scales and is therefore designed to act mostly in the zonal direction in the polar regions. For the diffusion acting on the wind field *before* the computation of the upstream position, simplified horizontal diffusion acting only in the zonal direction was consid-

ered. The results are shown in Fig. 3 and identified as the *1D diffusion* scheme. The results agree well with the full 2D diffusion scheme. The results to be presented later are all using the 1D diffusion. In fact, it was found that the inclusion of this diffusion did not make it necessary to have the usual 2D diffusion after the dynamical time step and it was then omitted in all our experiments.

5. Description of the 4DVAR configuration

To establish the final configuration of the 4DVAR assimilation, a number of experiments were carried out to determine the number of inner and outer loops in the incremental scheme and to characterize the simplified physics. The background error covariances used in the variational analysis (Gauthier et al. 1999b) act as a filter on the analysis increments. In 3DVAR, Laroche et al. (1999) have shown that in the incremental scheme, there is no gain in increasing the resolution of the variational analysis beyond what is necessary to represent the covariances. Given that the 4DVAR is using the same background error covariances as those of the 3DVAR analysis, the horizontal resolution of the 4DVAR has been kept the same as that of the 3DVAR, namely, at a triangular spectral truncation of T108 or 170-km horizontal resolution. Although this resolution is sufficient at the beginning of the assimilation window, small scales are developed during the 6-h tangent-linear and adjoint model integrations but can be resolved at this resolution. Moreover, as can be seen in Fig. 4, the energy spectra of the forecast error from 1 to 5 days indicates that 4DVAR is improving the large-scale components of the forecasts.

The presence of observation error correlation in satellite data limits the density of observations that can be assimilated. In Liu and Rabier (2003), if the observation error correlation is not taken into account in the assimilation, it is then preferable to thin the data so that the retained observations can be considered having uncorrelated observation error. At MSC, the satellite data are thinned at a resolution of 200 km or more, which is well resolved by the incremental model.

The trajectory used to define the tangent-linear and adjoint models requires an integration of the nonlinear model at full resolution and complete physical parameterizations. Trémolet (2004) presents a detailed study of different approaches to define the reference trajectory either by using an integration of the nonlinear model at a reduced resolution as first proposed by Rabier et al. (2000) or by interpolating the high-resolution trajectory to the lower resolution of the simplified model. In MSC's implementation of 4DVAR, the coarse resolution (170 km) and the high resolution (100 km) being rather close, for sake of simplicity, the tangent and adjoint model are defined with respect to a model trajectory generated by the nonlinear model with full physics being run at the same low resolution. Finally, the low-resolution increments need to be interpolated onto the high resolution of the operational model. MSC's GEM model being a gridpoint model, a cubic Lagrange scheme is used to interpolate the analysis increments from the low- to the high-resolution grid of the operational model.

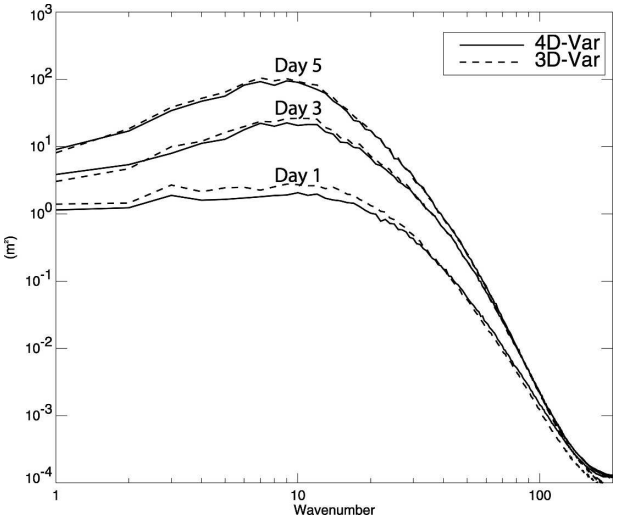

FIG. 4. Time evolution of the energy spectra of the forecast error of 3DVAR (dashed line) and 4DVAR (solid line) at 24, 72, and 120 h. Forecast error is measured here as the difference between the forecast and the verifying analysis. Each system is evaluated against its own analysis.

The simplified physics used within the inner loop includes simplified boundary layer dynamics (Laroche et al. 2002a), orographic blocking and stratiform condensation (Zadra et al. 2004), and convection (Mahfouf 2005). Two outer iterations were considered. The first inner loop was mostly adiabatic while the other included stratiform condensation, convection, and orographic blocking. The inclusion of stratiform condensation was motivated by the positive impact already noted by Mahfouf (1999). Finally, the formulation of a simplified deep convection proposed by Mahfouf (2005) is also included in the simplified physics. In Laroche et al. (2007), assimilation cycles over a period of one month were carried out to assess the impact of the different components introduced in the 4DVAR. These results show that, in our 4DVAR, the impact of the simplified physics improves the quality of the forecasts of 4DVAR by 7%. This will be further discussed in Laroche et al. (2007).

The total cost of 4DVAR is directly related to the number of iterations used to do the minimization. The minimization uses the quasi-Newton algorithm of Gilbert and LeMaréchal (1989). Experimentation was carried out with two outer iterations. The simplified physics is not adding a significant computer time (per iteration) and the total time is then proportional to the number of iterations. Figure 5 shows the results with 40 and 30 iterations in the two inner loops, the second including the full simplified physics. These results are compared with an experiment with 50 and 20 iterations in the inner loops. The final value of the cost function is evaluated after the final outer iteration where observations are compared against the high-resolution final analysis. Those final values are indicated in Fig. 5 by a black square (50–20 iterations) and a black circle (40–30 iterations). Both give nearly identical results. The figure also shows that complete convergence has not been reached in both cases. As the 3D/4DVAR cost function can be related to the a posteriori probability distribution, a reduction of the cost function implies that a more probable state has been found. So even if 3D/4DVAR has not converged, the resulting state is an improvement with respect to the background state. The norm of the gradient of the cost function in this particular case has been reduced by a factor of 10.

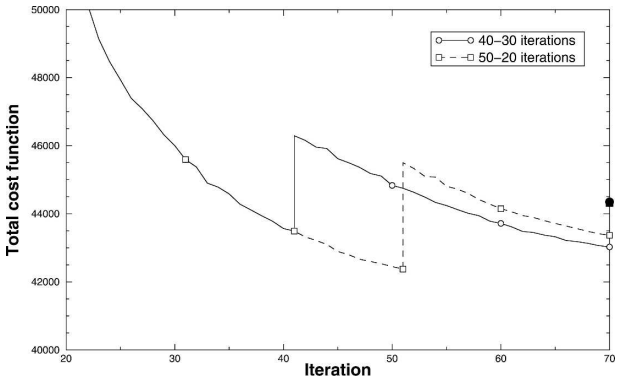

FIG. 5. Convergence of the minimization with different configurations for the inner loops with two outer iterations. The total number of iterations is 70 in both cases with 40 and 30 iterations in the first and second loop (solid line), respectively, and 50 and 20 iterations (dashed line). The black square and circle on the vertical axis at 70 iterations indicate the final values of both configurations after the final update in which the observations are compared to the final high-resolution analysis.

In an operational implementation, one is faced with the time constraint of having the analysis and forecast produced soon enough for the dissemination of products to the users. To achieve this objective, it has been necessary to configure the 4DVAR to meet this requirement without reducing the cutoff time for the reception of observations, keeping then the same number of observations. Therefore, after the first outer iteration, observations received during the 40 min or so that lasted the first minimization are added. This particularly benefits the TOVS data, at the end of the assimilation window. For the second minimization, the processing of all data is redone including the background check quality control and data thinning. The result from the first minimization provides a better initial point for the minimization and an estimate of the Hessian matrix obtained by the quasi-Newton minimizer. It is important to stress here that observations used in the first minimization may not be retained in the second pass but this makes the assimilation perfectly consistent. Finally, a larger number of iterations in the second loop is beneficial when observations are added at this stage. This is why the configuration with 40 and 30 iterations was retained instead of the one with 50 and 20 iterations.

Even though the 4DVAR takes approximately 1 h 20 min of wall clock time to run, the important element for the early cutoff run is the time it takes to do the second inner loop. In our case, it takes approximately 30 min and the analysis valid at time T can then be done at $T + 3$ h 10 min, and is available at $T + 4$ h. If the cutoff time was set with respect to the start of the 4DVAR analysis, the cutoff would then be set at 2 h 30 min. Figure 6 compares the volume of data that could be assimilated with a cutoff time of 2 h 30 min compared to the 3 h 10 min time used in the 4DVAR implementation. This shows that it would not be acceptable to fix the cutoff time at the time when the first inner loop is started. However, as this time constraint does not apply to the final analysis, it benefits from a very long cutoff time and it is this analysis that is used to produce the background state for the next analysis. This is an important point to remember because, when cycling 4DVAR, a significant part of the gain comes from the improvement in the background field that contains all the a priori information gained from past observations.

The tangent-linear and adjoint model integrations explain most of the total computing cost of 4DVAR. However, the wall clock time associated with this component can be reduced by the parallel implementation of the model by adding more CPUs. To achieve the goal of having completed the second inner loop in less than 20 min, the tangent-linear and adjoint model integrations were done using a total number of 40 CPUs of the IBM-p690 computer, which has a total number of 860 CPUs.

In 4DVAR, observations can be assimilated at the appropriate time over the whole assimilation window, here taken as 6 h, and the thinning algorithm has been

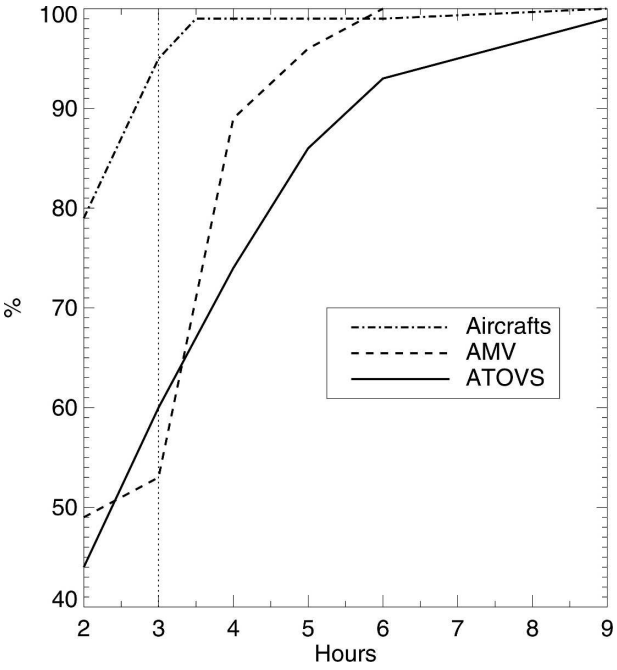

FIG. 6. Percentage of observations available as a function of cutoff time for the different observation types. The early cutoff (G1) occurs at $T + 3$ h 10 min while the reference analyses (G2) occur at $T + 6$ h at 0600 and 1800 UTC and $T + 9$ h at 0000 and 1200 UTC. At MSC, the forecasts are produced twice daily at 0000 and 1200 UTC.

altered as in the 4D screening scheme of Rabier et al. (2000). This has led to a significant increase in the volume of data ingested by the assimilation. Table 3 shows that this benefits the use of wind data from aircrafts and profilers in particular. In 3DVAR, it has been observed that wind data only within ±1 h 30 min from the analysis time can be kept: extending the window beyond leads to a degradation of the results. As expected satellite data, and particularly the atmospheric motion vectors (AMV), experienced a significant increase of admissible data. Otherwise, the observation operators of 4DVAR are the same as those of 3DVAR that act on profiled model states.

The experimentation and fine-tuning of all elements presented here have required quite an extensive level of experimentation because of the intimate relationship that now exists between the model and the whole assimilation process. More details on the analysis of the results obtained will be presented in a subsequent paper by Laroche et al. (2007). In the next section, the results obtained with the preoperational experimentation are presented. Those were obtained with the configuration of 4DVAR presented in this section.

To summarize, our implementation of the incremental 4DVAR comprises two outer iterations and the two

TABLE 3. Volumes of data assimilated in 4DVAR with respect to 3DVAR. The volume of added data is expressed relative to the number of data assimilated in 3DVAR.

Type	4DVAR	3DVAR	Volume of added data
Aircrafts	75 707	26 147	+190%
Radiosonde	66 605	66 603	~0%
SATWIND	82 160	41 604	+97%
ATOVS	71 517	46 832	+53%
GOES	3612	1979	+83%
Profilers	13 040	2196	+494%

inner loops comprise 40 iterations for the first one and the simplified physics only consider surface drag and vertical diffusion. After completion of the outer iteration, the second inner loop performs 30 more iterations with the stratiform condensation, convection, and orographic blocking added to the simplified physics. The simplified model operates at a resolution of 170 km (120×240 latitude–longitude Gaussian grid). The selection of observations is using the 4D screening approach of Rabier et al. (2000). However, given the higher computing cost of 4DVAR, it has been necessary to start the first inner loop before the cutoff time. Observations collected during the time taken to complete the first inner loop were added to the assimilation in the first outer iteration and the second inner loop. The background error statistics are the same as those used in 3DVAR (Gauthier et al. 1999b).

6. Impact of 4DVAR as implemented in March 2005: Results from complete two-month assimilation cycles

As is customary at MSC, extensive assimilation cycles were run and the resulting forecasts were evaluated for the boreal winter period (11 December 2003–11 February 2004) and the boreal summer period (15 July–15 September 2004). A summary of the results are now presented to show the impact of 4DVAR with respect to the previously operational 3DVAR analysis.

Figure 7 shows the verification of the analysis against radiosonde data during the summer period for the operational 3DVAR and 4DVAR. It shows that the 4DVAR analysis deviates more from the radiosonde data for all variables. However, in Fig. 8, the innovation error statistics show that 4DVAR is now in better agreement with the observations. This is a result that has often been obtained in 3DVAR when new data are introduced in the system. An analysis with a broader range of data often degrades the fit to a particular type of data. Another element is that with the radiosondes being available at synoptic time, the 3DVAR analysis

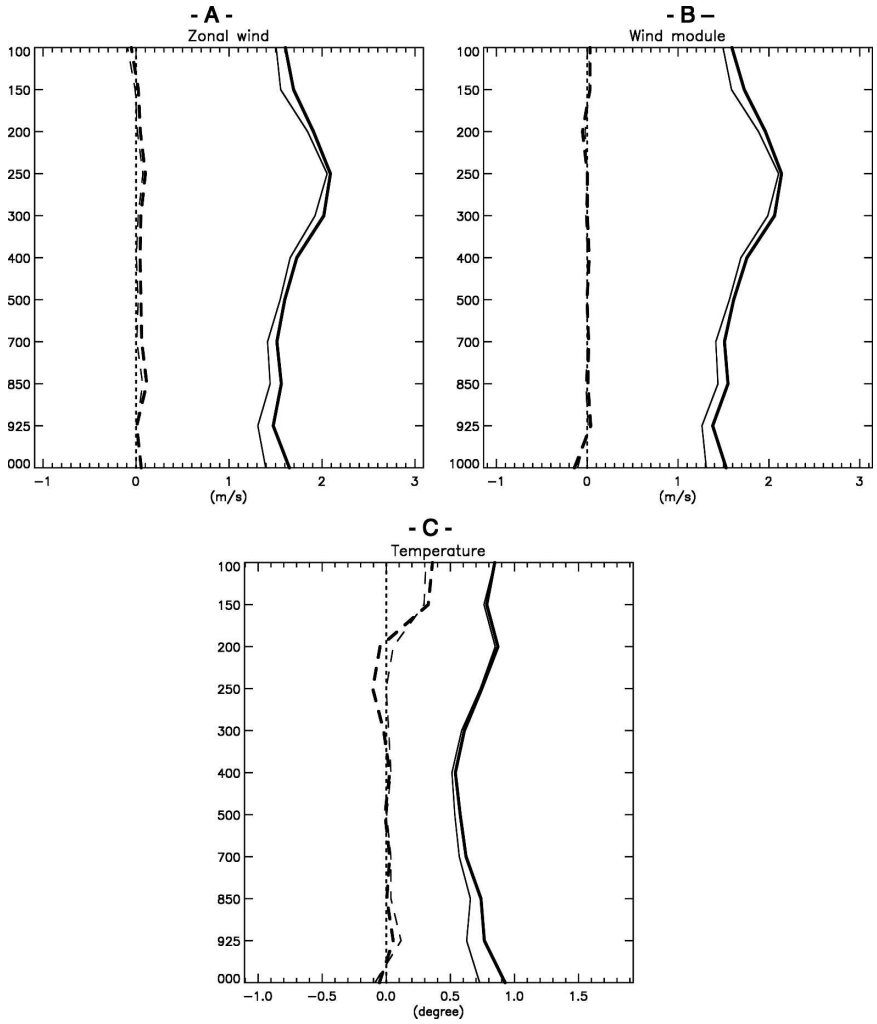

FIG. 7. Verification over the whole globe of the 3DVAR (thin line) and 4DVAR (solid line) analyses against radiosonde data for the summer period.

can more easily fit radiosonde data than the 4DVAR analysis, which needs to correct the initial conditions to fit data valid 3 h later. A 4DVAR analysis is then in fact the result of a 3-h forecast and this can contribute to the degradation of the fit to observations. However, the observation departure from the first guess is the true criterion as the impact of the analysis is evaluated against observations that have not been used by the analysis. Here the results show a positive impact with 4DVAR, particularly for winds.

The impact on the forecast at longer range has been independently evaluated for both the summer and winter periods. Verification against radiosonde data showed that the forecasts based on 4DVAR were improved over North America. The improvement was noted particularly over the west coast of North America at short ranges. This is illustrated by Fig. 9 showing the verification over this region and the whole Northern Hemisphere for the winter period. The improvement was noted over all regions and for nearly all objective criteria like rmse with respect to radiosonde or the verifying analysis. The verification against analyses over the Northern and Southern Hemisphere (Fig. 10) shows an overall improvement for both hemispheres but is more pronounced in the Southern Hemisphere where the analysis relies more on satellite data. This reflects the ability of 4DVAR to make better use of asynoptic data. This is also true in the Northern Hemisphere where 4DVAR is also making better use not only of satellite data but also of other asynoptic data from profilers and aircrafts. The interpretation of these results need to take into account the volume of data that has been significantly increased in 4DVAR. However, earlier experiments performed with exactly

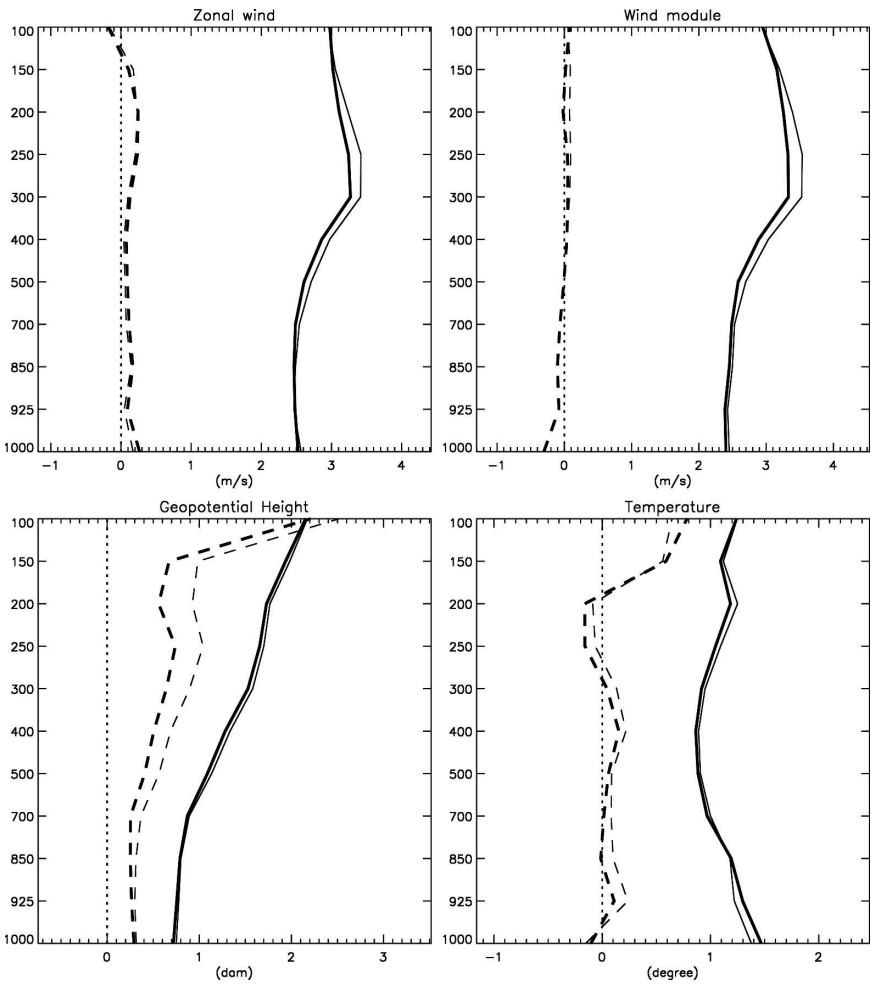

FIG. 8. Same as in Fig. 7, but for the 6-h forecast background error statistics. Verification of the geopotential is also shown here.

the same data (those used in the operational 3DVAR) have shown a positive impact. These results will be discussed in more details in Laroche et al. (2007).

In the regional forecast suite, the global analyses are used to initiate a short 3DVAR assimilation cycle with the higher-resolution regional model to establish dynamical balances proper to this model (Laroche et al. 1999). The GEM regional model configuration has a variable resolution with a 15-km horizontal resolution over North America. Significant differences also exist in the physical parameterizations: a full description can be found in Bélair et al. (2005). As shown by Gauthier and Thépaut (2001), a 4DVAR analysis is in better balance than what can be obtained from a sequential 3DVAR analysis. In particular, they showed that the 4DVAR analyses did not lead to a significant spinup in the precipitations in the early stages of the integrations. To assess if the 4DVAR analyses were palatable to the regional model, experiments were conducted in which the regional forecasts were initiated directly from the global 4DVAR analyses. Figure 11 compares the verifications of those forecasts against those from the operational system using its own 3DVAR analyses. As can be seen here, the 4DVAR analyses interpolated to the regional grid lead to better forecasts than those from the regional 3DVAR analyses: improvements in the synoptic scales of the 4DVAR analyses explain to a great extent the improvements noted in the regional forecasts. This also indicates that the 4DVAR global analyses have a dynamical balance that is more consistent with the dynamics of the regional model. However, this was not implemented as, in practice, the global analysis is only available *after* the regional forecast has been launched.

The 4DVAR assimilation was successfully implemented on 15 March 2005 in the operational suite of the Meteorological Service of Canada. Since then the system has been working without any problem and the

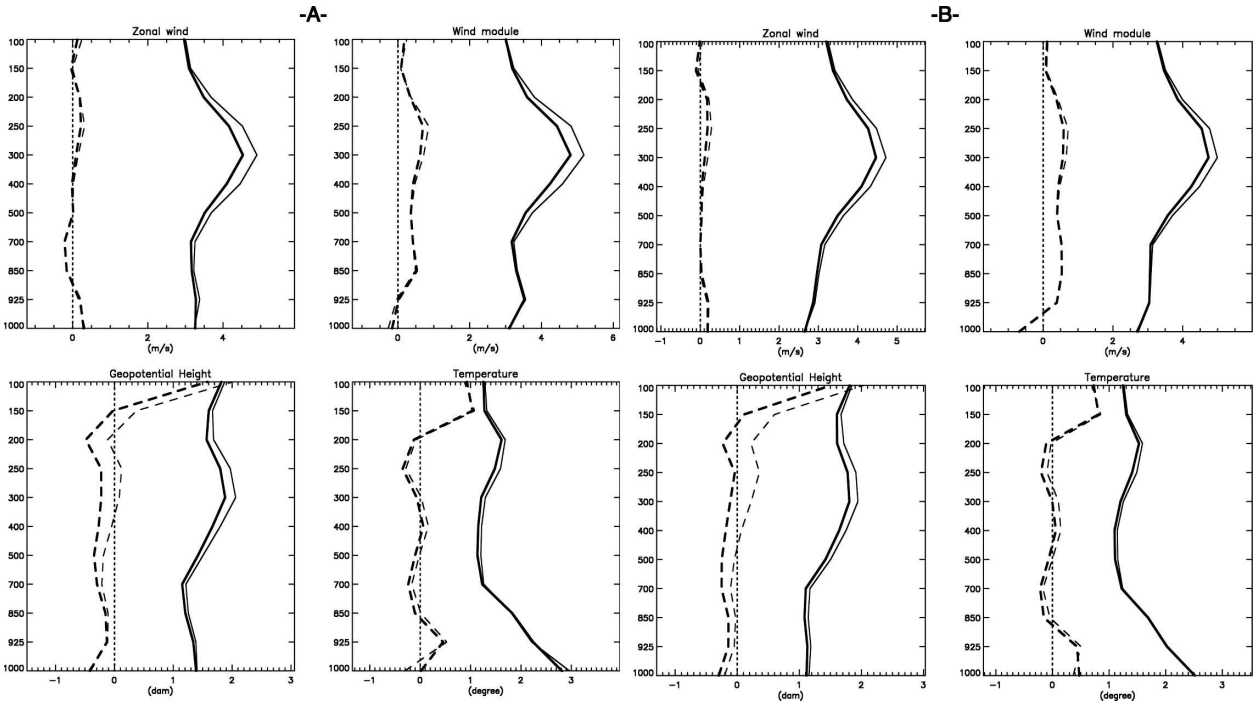

FIG. 9. Verification of 24-h forecasts issued from the operational 3DVAR (thin line) and the 4DVAR (solid line) analyses over (a) the western part of North America and (b) the whole Northern Hemisphere. The results are shown here for the winter period.

improvements noted in the preoperational experimentation were confirmed. The subjective evaluation of the operational forecasters was positive on all elements that are regularly examined. Their evaluation paid more attention to the forecasts over North America and Canada.

7. Conclusions

This paper gives a description of the 4DVAR data assimilation system as implemented at the Meteorological Service of Canada. The modularity of the architecture has permitted continuous development of both the observations and background term aspects in 3DVAR on the one hand and of the global model on the other hand. This has made it easier to manage the development work done by several people. In the end, the system succeeded to meet all the operational time constraints while at the same time improving the quality of MSC's analysis and forecast products. The treatment of observations (i.e., observation operators) and error statistics in 4DVAR were kept identical to the previous 3DVAR assimilation system except for the 4D thinning of data that has led to a significant increase in the volume of assimilated data. As suggested by Gauthier and Thépaut (2001), 4DVAR analyses do not lead to internal dynamical imbalances that can be detrimental to the forecasts. In MSC's implementation, there is no penalty term to control the emergence of gravity waves. It should be said, however, that the forecasts are nevertheless using the digital filter finalization technique of Fillion et al. (1995).

The development and implementation of 4DVAR at the Meteorological Service of Canada is the result of many years of effort that began in 1994. At that time, a new gridpoint model was being developed (Côté et al. 1998a) taking into account the specific needs of data assimilation by including the necessary effort to have its tangent-linear and adjoint models (Tanguay and Polavarapu 1999). At the same time, the development of the 3DVAR was initiated and implemented in 1997 in the global (Gauthier et al. 1999a) and regional (Laroche et al. 1999) operational suites. The direct assimilation of satellite radiances constituted the third line of work that was started at about the same time. Satellite radiances have been used in the 3DVAR since September 2000. In this paper, the supplementary work required to obtain a 4DVAR by coupling the BGOBS unit to the TLM/adjoint models has been described. This approach has made it possible to pursue the independent development of the 3DVAR and the model components. When the experimentation with 4DVAR began, it then benefited from all observations available to the 3DVAR and of the numerous changes brought to the

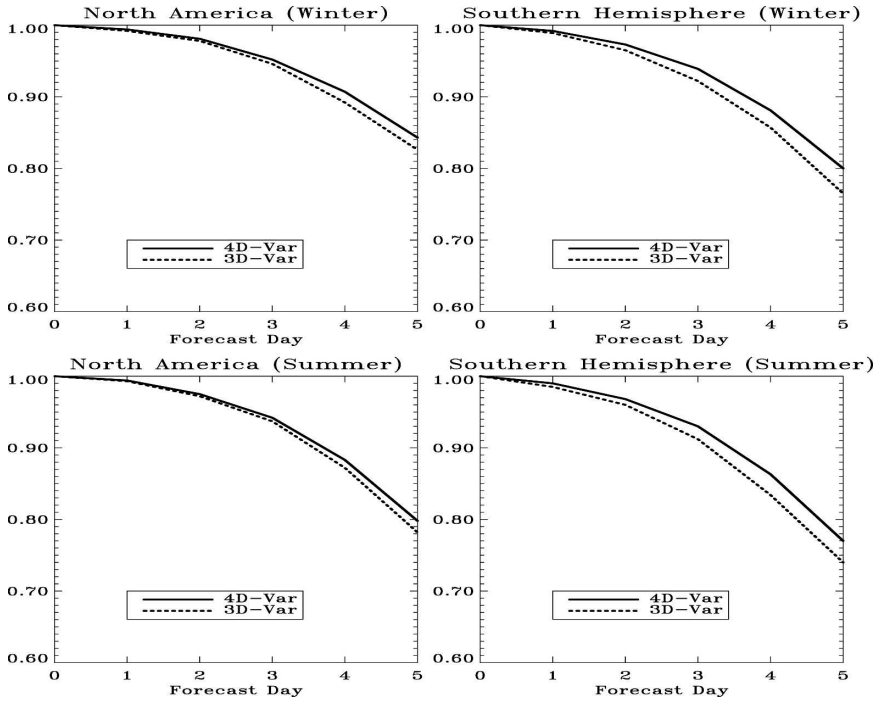

FIG. 10. Anomaly correlation scores for 4DVAR (solid line) and 3DVAR (dotted line) for the summer (15 Jul–15 Sep 2004) and winter (11 Dec 2003–11 Feb 2004) periods over North America and the Southern Hemisphere.

model itself. This included computing, numerical, and physical changes required to improve the operational forecasts and maintain computational efficiency. This also included modifications due to a significant change in the computing architecture when MSC changed its NEC SX-6 computer for an IBM-p690.

All in all, the implementation of 4DVAR, although arduous, was successful and is now providing the proper framework to add more satellite data to the assimilation system. This includes for instance the SSM/I (Deblonde 1999), the Atmospheric Infrared Sounder (AIRS; Garand et al. 2005) radiances, and GPS radio occultations (Aparicio and Deblonde 2007). These observations should be added to the operational system in the near future. However, the limited computer resources have required that work be concentrated on upgrading the global model by increasing its resolution from the current 100 to 35 km. This version will also include a complete revision of the physical parameterizations (Bélair et al. 2005). It is anticipated that this version should become operational in 2006.

The modular nature of our system is also being adapted to obtain a 4DVAR for the limited-area version of the GEM model. Preliminary experimentation is also exploring the inclusion of stratospheric chemistry in a fully coupled atmosphere–chemistry system to study the two-way interactions between dynamics and atmospheric chemistry (Ménard 2005).

Acknowledgments. Several people contributed to this implementation to resolve many elements that were needed. Michel Desgagné, Vivian Lee, and Michel Valin contributed to the development and optimization of the tangent-linear and adjoint models, and also to the design of the complex launching scripts of the 4DVAR. We also acknowledge the significant contribution from the Data Assimilation Development Section of the Meteorological Service of Canada (MSC) who have addressed all the observation processing elements, including reception, monitoring, and archiving. The contributions of Gilles Verner, José Garcia, Lorraine Veillette, Réal Sarrazin, Jacques Hallé, Pierre Koclas, Nicolas Wagneur, Nils Ek, and Judy St. James have been particularly important. MSC also provided technical support for the implementation. At the final stages, the operational forecasters provided valuable feedback that helped us iron out some remaining glitches. Their subjective evaluation helped us to ensure that all aspects of the analyses and forecasts used by operations were indeed meeting or exceeding their expectations. We would like to acknowledge particularly the involvement of Alan Rahill and Suzanne Roy. Comments on

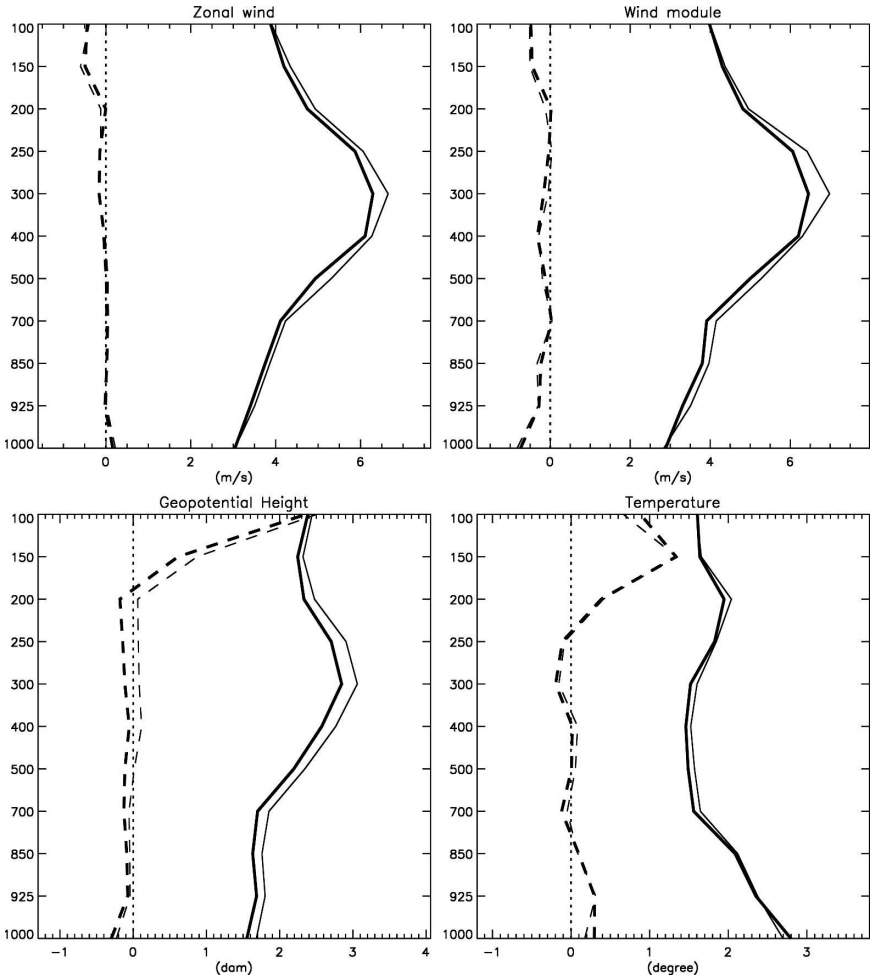

FIG. 11. Verification of regional 48-h forecasts against radiosondes obtained from the regional operational system using its own 3DVAR analysis (thin line) and regional forecasts initiated directly from the 4DVAR global analyses (thick line). The results shown are based on 24 winter cases from the winter period extending from 11 Dec 2003 to 11 Feb 2004.

the manuscript by Erik Andersson, Stephen Cohn, and an anonymous reviewer are also acknowledged.

The operational 4DVAR of MSC uses the M1QN3 minimization code provided to us by Jean-Charles Gilbert of the Institut National de Recherche en Informatique et en Automatique (INRIA).

APPENDIX

Vertical Interpolation for Observation Operators

The fact that the observation operators characterize the vertical levels only in terms of pressure meant that it sufficed to modify the computation of these pressures given as input to the observation operators. To make this idea more precise, consider a single observation of temperature defined at $p = p_{\text{obs}}$. The vertical coordinate η of the model being defined as

$$\eta_k(p) = \frac{(p - p_T)}{(p_s - p_T)},$$

so that $p_k < p_{\text{obs}} \leq p_{k+1}$ with $p_k = p_T + \eta_k (p_s - p_T)$ with $p_T = 10$ hPa being the prescribed pressure at the model's lid (Côté et al. 1998a). Linear interpolation is performed in terms of ln p and the observation operator is then

$$\mathbf{H}(T_k, T_{k+1}, p_s) \equiv \alpha_k(p_s)T_k + [1 - \alpha_k(p_s)]T_{k+1},$$

where

$$\alpha_k(p_s) = \frac{\ln p_{\text{obs}} - \ln p_k}{\ln p_{k+1} - \ln p_k}.$$

The dependence on p_s makes $\mathbf{H}$ nonlinear and its tangent linear then reads as

$$\frac{\partial \mathbf{H}}{\partial \mathbf{X}}(T_k, T_{k+1}, p_s) = \left[\alpha_k \quad (1 - \alpha_k) \quad (T_{k+1} - T_k) \frac{\partial \alpha_k}{\partial p_s} \right],$$

where

$$\frac{\partial \alpha_k}{\partial p_s} = - \frac{1}{\ln p_{k+1} - \ln p_k} \left[(1 - \alpha_k) \frac{\eta_k}{p_k} + \alpha_k \frac{\eta_{k+1}}{p_{k+1}} \right].$$

This implies that changes in p_s need to be taken into account when the vertical coordinate depends on surface pressure and the impact depends on the lapse rate of the *reference state*. This simple example shows that the observation operator can be defined provided the pressure values associated with each model level is given at each observation location. This makes it very easy to change the vertical coordinate in the 3DVAR. This has been found to be useful for the introduction of a new hybrid coordinate in the GEM model or to adapt it for the midatmospheric assimilation system of Polavarapu et al. (2005).

REFERENCES

Andersson, E., and H. Järvinen, 1999: Variational quality control. *Quart. J. Roy. Meteor. Soc.,* **125,** 697–722.

Aparicio, J., and G. Deblonde, 2007: Evaluation of GPS radio occultation assimilation. *Mon. Wea. Rev.,* in press.

Bélair, S., J. Mailhot, C. Girard, and P. Vaillancourt, 2005: Boundary layer and shallow cumulus clouds in a medium-range forecast of a large-scale weather system. *Mon. Wea. Rev.,* **133,** 1938–1960.

Chouinard, C., C. Charette, J. Hallé, P. Gauthier, J. Morneau, and R. Sarrazin, 2001: The Canadian 3D-Var analysis scheme on model vertical coordinate. Preprints, *14th Conf. on Numerical Weather Prediction,* Fort Lauderdale, FL, Amer. Meteor. Soc., 14–18.

Côté, J., S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. N. Staniforth, 1998a: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.,* **126,** 1373–1395.

——, J. G. Desmarais, S. Gravel, A. Méthot, A. Patoine, M. Roch, and A. N. Staniforth, 1998b: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part II: Results. *Mon. Wea. Rev.,* **126,** 1397–1418.

Courtier, P., 1997: Dual formulation of four-dimensional variational data assimilation. *Quart. J. Roy. Meteor. Soc.,* **123,** 2449–2461.

——, J. N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var, using an incremental approach. *Quart. J. Roy. Meteor. Soc.,* **120,** 1367–1387.

Deblonde, G., 1999: Variational assimilation of SSM/I total precipitable water retrievals in the CMC analysis system. *Mon. Wea. Rev.,* **127,** 1458–1476.

Desgagné, M., S. Thomas, and M. Valin, 2000: Performance of MC2 and the ECMWF IFS forecast model on the Fujitsu VPP700 and NEC SX 4M. *Sci. Program.,* **8,** 23–30.

Durran, D. R., 1999: *Numerical Methods for Wave Equations in Geophysical Fluid Dynamics.* Springer Verlag, 465 pp.

Eyre, J. R., 1989: Inversion of cloudy satellite sounding radiances by nonlinear optimal estimation. I: Theory and simulation of TOVS. *Quart. J. Roy. Meteor. Soc.,* **115,** 1001–1026.

Fillion, L., H. L. Mitchell, H. R. Ritchie, and A. N. Staniforth, 1995: The impact of a digital filter finalization technique in a global data assimilation system. *Tellus,* **47A,** 304–323.

Garand, L., A. Beaulne, and N. Wagneur, 2005: Assimilation of AIRS hyperspectral radiances at the Meteorological Service of Canada. Preprints, *14th Conf. on Satellite Meteorology and Oceanography,* Atlanta, GA, Amer. Meteor. Soc., CD-ROM, P5.13.

Gauthier, P., 2003: Operational implementation of variational data assimilation. *Data Assimilation for the Earth System,* R. Swinbank, V. Shutyaev, and W. A. Lahoz, Eds., NATO Science Series, IV: Earth and Environmental Sciences, Vol. 26, Kluwer Academic, 167–176.

——, and J.-N. Thépaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.,* **129,** 2089–2102.

——, C. Charette, L. Fillion, P. Koclas, and S. Laroche, 1999a: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part I: The global analysis. *Atmos.–Ocean,* **37,** 103–156.

——, M. Buehner, and L. Fillion, 1999b: Background-error statistics modelling in a 3D variational data assimilation scheme: Estimation and impact on the analyses. *Proc. ECMWF Workshop on Diagnosis of Data Assimilation Systems,* Reading, United Kingdom, ECMWF, 131–145.

——, C. Chouinard, and B. Brasnett, 2003: Quality control: Methodology and applications. *Data Assimilation for the Earth System,* R. Swinbank, V. Shutyaev, and W. A. Lahoz, Eds., NATO Science Series, IV: Earth and Environmental Sciences, Vol. 26, Kluwer Academic, 177–187.

Gilbert, J. C., and C. LeMaréchal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Program.,* **45,** 407–435.

Ingleby, N. B., and A. C. Lorenc, 1993: Bayesian quality control using multivariate normal distributions. *Quart. J. Roy. Meteor. Soc.,* **119,** 1195–1225.

Lagarde, T., A. Piacentini, and O. Thual, 2001: A new representation of data assimilation methods: The PALM flowcharting approach. *Quart. J. Roy. Meteor. Soc.,* **127,** 189–207.

Laroche, S., and P. Gauthier, 1998: A validation of the incremental formulation of 4D variational data assimilation in a nonlinear barotropic flow. *Tellus,* **50A,** 557–572.

——, ——, J. St. James, and J. Morneau, 1999: Implementation of a 3D variational data assimilation system at the Canadian Meteorological Centre. Part II: The regional analysis. *Atmos.–Ocean,* **37,** 281–307.

——, M. Tanguay, and Y. Delage, 2002a: Linearization of a simplified planetary boundary layer parameterization. *Mon. Wea. Rev.,* **130,** 2074–2087.

——, ——, A. Zadra, and J. Morneau, 2002b: Use of adjoint sensitivity analysis to diagnose the CMC global analysis performance: A case study. *Atmos.–Ocean,* **40,** 423–443.

——, P. Gauthier, M. Tanguay, S. Pellerin, and J. Morneau, 2007: Impact of the different components of 4DVAR on the global forecast system of the Meteorological Service of Canada. *Mon. Wea. Rev.,* **135,** 2355–2364.

Liu, Z. Q., and F. Rabier, 2003: The potential of high-density observations for numerical weather prediction: A study with simulated observations. *Quart. J. Roy. Meteor. Soc.,* **129,** 3013–3035.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.,* **112,** 1177–1194.

Mahfouf, J.-F., 1999: Influence of physical processes on the tangent-linear approximation. *Tellus,* **51A,** 147–166.

——, 2005: Linearization of a simple moist convection scheme for large-scale NWP models. *Mon. Wea. Rev.,* **133,** 1655–1670.

Ménard, R., 2005: Advances in atmospheric chemistry and dynamics research by development of coupled chemistry-dynamics data assimilation. ESTEC Contract 18560/04/NL/FF, 116 pp.

Pellerin, S., P. Gauthier, M. Tanguay, S. Laroche, and J. Morneau, 2005: A modular implementation of the operational 4D-Var data assimilation at the Meteorological Service of Canada. *Proc. Fourth WMO Symp. on Assimilation of Observations in Meteorology and Oceanography,* Prague, Czech Republic, WMO, WWRP 9, WMO-TD 1316, CD-ROM.

Polavarapu, S., and M. Tanguay, 1998: Linearizing iterative processes for four-dimensional data assimilation schemes. *Quart. J. Roy. Meteor. Soc.,* **124,** 1715–1742.

——, ——, R. Ménard, and A. Staniforth, 1996: The tangent linear model for semi-Lagrangian schemes: Linearizing the process of interpolation. *Tellus,* **48A,** 74–95.

——, S. Ren, Y. Rochon, D. Sankey, N. Ek, J. Koshyk, and D. Tarasick, 2005: Data assimilation with the Canadian Middle Atmosphere Model. *Atmos.–Ocean,* **44,** 77–100.

Pudykiewicz, J., and A. Staniforth, 1984: Some properties and comparative performance of the semi-Lagrangian method of Robert in the solution of the advection-diffusion equation. *Atmos.–Ocean,* **22,** 283–308.

Qaddouri, A., J. Côté, and M. Valin, 1999: A parallel direct 3D elliptic solver. *High Performance Computing Systems and Applications,* A. Pollard, D. Mewhort, and D. Weaver, Eds., Kluwer Academic, 429–442.

Rabier, F., H. Järvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four dimensional variational assimilation. Part I: Experimental results with simplified physics. *Quart. J. Roy. Meteor. Soc.,* **126,** 1143–1170.

Smolarkiewicz, P., and J. Pudykiewicz, 1992: A class of semi-Lagrangian approximations for fluids. *J. Atmos. Sci.,* **49,** 2082–2096.

Talagrand, O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. I—Theory. *Quart. J. Roy. Meteor. Soc.,* **113,** 1311–1328.

Tanguay, M., and S. Polavarapu, 1999: The adjoint of the semi-Lagrangian treatment of the passive tracer equation. *Mon. Wea. Rev.,* **127,** 551–564.

——, ——, and P. Gauthier, 1997: Temporal accumulation of first-order linearization error for semi-Lagrangian passive advection. *Mon. Wea. Rev.,* **125,** 1296–1311.

Trémolet, Y., 2004: Diagnostics of linear and incremental approximations in 4D-Var. *Quart. J. Roy. Meteor. Soc.,* **130,** 2233–2251.

Zadra, A., M. Buehner, S. Laroche, and J.-F. Mahfouf, 2004: Impact of the GEM model simplified physics on the extratropical singular vectors. *Quart. J. Roy. Meteor. Soc.,* **130,** 2541–2569.